Normalisation.doc G00376315 Elizabeth Daly

* As per the lectures, normalisation is the process of organising the columns and tables of a relational database to minimise data redundancy. Normalisation should also eliminate insert, delete, and update anomalies, and simplify queries.
* In this database there is a lot of data redundancy because each time a student enrols in a new module the table gets a new row, and those two rows contain a lot of the same information. For example, rows 3 and 4 contain much of the same information (studentID, studentName, and dob attributes) while only the moduleID and moduleName columns have different values. So, the design is not good because some data is repeated within certain single cells.
* The primary key (PK) must uniquely identify each record/row in a table. It can consist of multiple fields, as it does here because the primary key is a composite primary key consisting of studentID and moduleID columns. For example, rows 5 and 6 are uniquely identified because row 5 has PK (studentID=4 and moduleID=101) while row 6 has PK (studentID=4 and moduleID=102). The PK is good so the design is good from this point of view.
* Here we are told that a student can enrol before deciding which modules to take, so initially, the moduleID column could be empty/unknown. That would also be the case if a module was not available in a given year. These situations would correspond to one of the PKs having a NULL value, and PKs cannot contain NULL values. The design is not good for this reason.
* If a module was not taken by any student, then the studentID, studentName, and dob columns would be empty. That means that these fields would contain NULL; studentID is a PK and cannot have a NULL value. The design is not good for this reason.
* Proposed alternative design in Innovation.doc